

Interaction between Truth and Belief as the key to entropy and other quantities of statistical physics

Flemming Topsøe

University of Copenhagen, Department of Mathematical Sciences
Universitetsparken 5, 2100 Copenhagen, Denmark
topsoe@math.ku.dk

Abstract

The notion of *entropy* penetrates much of science. A key feature of the all-important notion of *Boltzmann-Gibbs-Shannon entropy* is its *extensivity* (additivity over independent subsystems). However, there is a need for other quantities. In statistical physics a parameterized family of *non-extensive* entropy measures, now mainly known under the name of *Tsallis-entropies*, have received much attention but also been met with criticism due mainly to a lack of convincing interpretations.

Based on the hypothesis that interaction between *truth*, held by “Nature”, and *belief*, as expressed by man, may take place, classical- as well as non-classical measures of entropy and other essential quantities are derived. The approach aims at providing a genuine interpretation, rather than relying either on analogies based on formal mathematical manipulations or else – more fruitfully, but not satisfactory – on axiomatic characterizations.

1 Contemplation

Let us apply a philosophical approach and put ourselves in the shoes of the physicist, planning to set-up *experiments* and to engage in associated *observations*. He might argue as follows:

1: As an expression of my *beliefs* concerning phenomena I plan to observe, I shall assign numbers in $[0, 1]$, typically denoted by the letter y , to *events* associated with the phenomena. Certain, to me unknown numbers, likewise in $[0, 1]$, and likewise associated with events, express the essence of the phenomena, and do not depend on my interference. They are referred to as *truth-assignments* and are, typically, denoted by the letter x .

2: Any event I may observe entails a certain *effort* on my part. This effort I shall also refer to as *individual complexity* – “individual”, because it is associated with each individual event I could encounter. Before setting up *experiments*, I should determine the effort I am willing to or have to devote to any event I may be faced with. This should depend only on the assigned belief-value y , and is denoted $\kappa(y)$. The function κ , defined on $[0, 1]$ and with values in $[0, \infty]$, I refer to as the *coder*. As 1 represents *certainly*, I insist that $\kappa(1) = 0$. Further, I assume that κ is smooth in a technical sense, say continuous on $[0, 1]$ and continuously differentiable and finite valued on $]0, 1]$. Finally, as I do not want to distinguish between coders that only differ by a scalar factor, I will introduce an assumption of *normalization* in order to pick out a canonical representative of the possible coders. As $\kappa(1) = 0$ and as I do not want to assume that $\kappa(0)$ is finite, I choose to impose the condition $\kappa'(1) = -1$ for the stated purpose.

3 To determine the coder, I must know the basic characteristics of the *world* I operate in. I choose to focus primarily on a concept of *interaction between truth and belief*.

4 I shall model this interaction by a function π defined on the product set $[0, 1] \times [0, 1]$ and taking values in $]-\infty, \infty]$. The idea is that $\pi(x, y)$ represents the *force* by which the world presents an event to me in case the truth-assignment is x and my belief in the event is y . On the technical side, I better assume that π is continuous on its domain and continuously differentiable and finite-valued on $[0, 1] \times]0, 1]$.

5 I consider the *classical world* to be a world of “*no interaction*”, i.e. $\pi(x, y) = x$ for all (x, y) . I must be prepared for other forms of interaction, but will always assume that the interaction is *sound*, i.e. that $\pi(x, x) = x$ for all $x \in [0, 1]$. Stronger conditions should be considered and in this connection, it appears sensible to impose conditions of *consistency*: I will call the interaction *weakly consistent* if $\sum_{i \in \mathbb{A}} \pi(x_i, y_i) = 1$ for any finite *alphabet* \mathbb{A} and any truth-assignment $x = (x_i)_{i \in \mathbb{A}}$ and belief-assignment $y = (y_i)_{i \in \mathbb{A}}$,

both assumed to be probability distributions over \mathbb{A} . If, with the same conditions on x and y , it can be concluded that $(\pi_i)_{i \in \mathbb{A}} = (\pi(x_i, y_i))_{i \in \mathbb{A}}$ is in fact a probability distribution, just as x and y , I will say that π is *strongly consistent*.

6 To enable *observations* from the world, I must *configure* all available resources such as *observation- and measuring devices*. The resulting *configuration* will enable me to perform *experiments*, i.e. to study individual *situations* from the world which have my interest.

7 Before actual observations are performed, I must identify the various possible *basic events* (or *pure states*, as some may prefer), which I could encounter. I shall characterize them by an *index*, typically i , intended to have *semantic significance*. The set of possible basic events is the *alphabet* pertaining to the situation, call it \mathbb{A} . The actual naming of members of \mathbb{A} , the *semiotic assignment*, should catalyze semantic awareness and facilitate technical handling.

8 I will apply a *principle of separability* and consider my *total effort* related to observations from the configured situation to be the sum of individual efforts associated with the basic events. In so doing, I must weigh each contribution according to the force with which I will experience the associated basic event. The total effort I also refer to as *total complexity* or simply *complexity* and thus find that complexity is the weighted sum of individual complexities:

$$\Phi(x, y) = \sum_{i \in \mathbb{A}} \pi(x_i, y_i) \kappa(y_i). \quad (1)$$

Here, $x = (x_i)_{i \in \mathbb{A}}$ and $y = (y_i)_{i \in \mathbb{A}}$ are, respectively the truth-assignments and the belief-assignments associated with the various basic events.

9 I will attempt to minimize complexity and shall appeal to the principle that the smallest value for complexity is obtained when *belief matches truth*. As

$$\sum_{i \in \mathbb{A}} \pi(x_i, y_i) \kappa(y_i) - \sum_{i \in \mathbb{A}} x_i \kappa(x_i) \quad (2)$$

represents my *frustration*, the principle says that frustration is the least, in fact disappears, when $y_i = x_i$ for all $i \in \mathbb{A}$.

Given $x = (x_i)_{i \in \mathbb{A}}$, minimal complexity is what I will aim at. It is an important quantity. In anticipation, I will call it *entropy* and denote it by

the letter H^1 :

$$H(x) = \inf_{y=(y_i)_{i \in \mathbb{A}}} \Phi(x, y) = \sum_{i \in \mathbb{A}} x_i \kappa(x_i). \quad (3)$$

The quantity (2) too appears important. It is tempting to call it “*frustration*” but, again in anticipation, I better call it *divergence*. I shall denote it by the letter D :

$$D(x, y) = \Phi(x, y) - H(x). \quad (4)$$

2 Conclusion

Theorem 1. *With assumptions and definitions as introduced above, assuming only that the interaction is weakly consistent, the number $q = \pi(1, 0)$ must be non-negative and, to each $q \in [0, \infty[$, there is only one pair of interaction and coder which fulfill the conditions imposed. These functions, denoted π_q and κ_q , are determined by the formulas*

$$\pi_q(x, y) = qx + (1 - q)y, \quad (5)$$

$$\kappa_q(y) = \ln_q \frac{1}{y}, \quad (6)$$

where the q -logarithm is given by

$$\ln_q x = \begin{cases} \ln x & \text{if } q = 1, \\ \frac{x^{1-q} - 1}{1-q} & \text{if } q \neq 1. \end{cases} \quad (7)$$

Note that strong consistency holds if and only if $0 \leq q \leq 1$.

The accompanying quantities, complexity, entropy and divergence are denoted Φ_q , H_q and D_q , respectively, and given through (1), (3) and (4), i.e.

$$\Phi_q(x, y) = \sum_{i \in \mathbb{A}} \pi_q(x_i, y_i) \kappa_q(y_i), \quad (8)$$

$$H_q(x) = \sum_{i \in \mathbb{A}} x_i \kappa_q(x_i) \quad (9)$$

$$D_q(x, y) = \sum_{i \in \mathbb{A}} \left(\pi_q(x_i, y_i) \kappa_q(y_i) - x_i \kappa_q(x_i) \right). \quad (10)$$

¹In order to allow a singular case – the case $q = 0$ of Theorem 1 below – to fit into the framework, the infimum in (3) should be restricted to run over probability distributions y with a support which contains the support of x .

In (9) we recognize the family of *Tsallis entropies*, cf. Tsallis [6].

Regarding the proof of Theorem 1, we shall here only give a brief indication: The formula (5) is readily derived from the assumption of weak consistency. Then, the only possible form for the coder, (6), is derived from pretty standard variational arguments. The final step of the proof, that with (5) and (6) the variational principle does indeed hold, follows by observing the close tie to entropy- and divergence- measures as derived by an approach due to Bregman, cf. the recent papers [5] and [4] and references there.

3 Hints to the literature

Regarding the formula (9) for entropy and its significance, we note that it first appeared in the mathematical literature in Havrda and Charvát [1], that it then appeared in the physical literature in Lindhard and Nielsen [3], and in Lindhard [2], and that it was efficiently promoted in the paper by Tsallis [6] which triggered much research in the physical community as also witnessed by the many entries in the database pointed to under [6].

4 Formal publication

The present manuscript, posted on the arXiv server, is an announcement, prior to formal publication. A manuscript with a comprehensive discussion, with a full proof of Theorem 1 and with some further results will be worked out soon and submitted to the electronic journal “Entropy”.

Further discussion of the considerations presented as well as suggestions of concrete mechanisms behind the concept of interaction are among obvious issues to look closer into.

References

- [1] J. Havrda and F. Charvát. Quantification method of classification processes. Concept of structural α -entropy. *Kybernetika*, 3:30–35, 1967.
- [2] J. Lindhard. On the Theory of Measurement and its Consequences in Statistical Dynamics. *Mat. Fys. Medd. Dan. Vid. Selsk.*, 39(1):1–39, 1974.

- [3] J. Lindhard and V. Nielsen. Studies in Statistical Dynamics. *Mat. Fys. Medd. Dan. Vid. Selsk.*, 38(9):1–42, 1971.
- [4] J. Naudts. Generalised exponential families and associated entropy functions. arXiv:0803.0104v1[math-ph], 2008.
- [5] F. Topsøe. Exponential Families and MaxEnt Calculations for Entropy Measures of Statistical Physics, in volume 965 of *AIP Conference Proceedings*, pages 104–113, 2007.
- [6] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Physics*, 52:479–487, 1988. See <http://tsallis.cat.cbpf.br/biblio.htm> for an updated bibliography.